Lesson 18: Comparing Populations Using Samples

Goals

- Calculate the difference between the mean or median of two samples from different populations, and express it as a multiple of the MAD or IQR.
- Interpret a pair of box plots, including the amount of visual overlap between the two distributions.
- Justify (orally and in writing) whether there is likely to be a meaningful difference between two populations, based on a sample from each population.

Learning Targets

- I can calculate the difference between two medians as a multiple of the interquartile range.
- I can determine whether there is a meaningful difference between two populations based on a sample from each population.

Lesson Narrative

In previous lessons, students examined the distributions of two entire populations to decide whether or not they were very different. In this lesson, students use samples to make comparative inferences about populations.

Students see that if samples of two different populations have only a small difference between their measures of center (relative to their variability), then we cannot say that there is a meaningful difference between the measures of center of the populations (MP2). Due to sampling variability, it is possible that the two populations may not be very different. However, if samples from two different populations have a large difference between their measures of center (relative to their variability), then we can say that there is likely to be a meaningful difference between the measures of center of the two populations.

Alignments

Building On

- 6.SP.B.5.c: Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

Addressing

- 7.SP.B.3: Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the

variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.

- 7.SP.B.4: Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.

Building Towards

- 7.SP.B: Draw informal comparative inferences about two populations.

Instructional Routines

- MLR1: Stronger and Clearer Each Time
- MLR5: Co-Craft Questions
- MLR7: Compare and Connect
- Think Pair Share

> ### Student Learning Goals
>
> Let's compare different populations using samples.

18.1 Same Mean? Same MAD?

Warm Up: 5 minutes
This warm-up reminds students of the meanings of mean and MAD by comparing two sets of data with similar but different values and asking whether they will have the same means or MADs or both.

Building On

- 6.SP.B.5.c

Building Towards

- 7.SP.B

Launch

Explain to students that the pairs of data sets are: A and B, X and Y, and P and Q.

Anticipated Misconceptions

For students who have a difficult time starting without calculating, help them to compare the values in the ones place for the first and third pairs of data.

> ### Student Task Statement
>
> Without calculating, tell whether each pair of data sets have the same mean and whether they have the same mean absolute deviation.

1. set A 1 3 3 5 6 8 10 14

   set B 21 23 23 25 26 28 30 34

2. set X 1 2 3 4 5

   set Y 1 2 3 4 5 6

3. set P 47 53 58 62

   set Q 37 43 68 72

Student Response

1. Data sets A and B have different means, but the same MADs.
2. Data sets X and Y have different means and different MADs.
3. Data sets P and Q have the same means, but different MADs.

Activity Synthesis

The purpose of the discussion is to bring out methods students used to notice whether the pairs of data sets had the same mean or MAD or both.

Poll the class for each pair of data sets as to whether they had the same mean, MAD, both, or neither.

After students have had a chance to register their vote, ask some students to explain their reasoning for their answer.

18.2 With a Heavy Load

10 minutes

In a previous lesson, students compared heights of two teams of people when the entire populations were known. In this activity, students only have access to data from a sample of each population and are asked to determine if the populations are different based on the sample.

Students construct informal arguments to explain why the different samples come from populations that are meaningfully different or not (MP3).

Addressing

- 7.SP.B.3
- 7.SP.B.4

Instructional Routines

- MLR5: Co-Craft Questions
- Think Pair Share

Launch

Arrange students in groups of 2. Allow students 3–5 minutes quiet work time followed by partner and whole-class discussions.

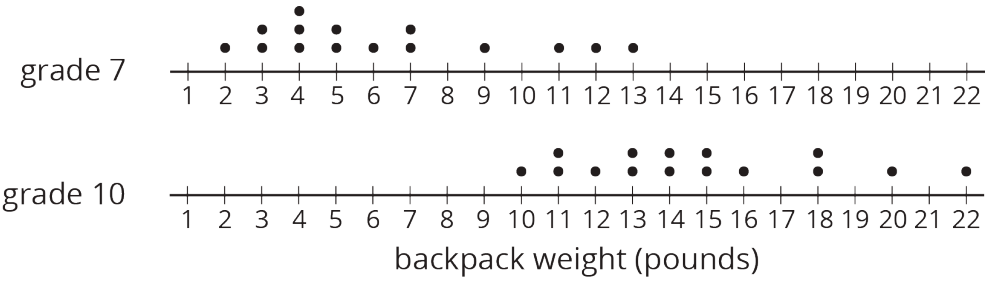
Access for English Language Learners

Writing, Conversing: MLR5 Co-Craft Questions. Display the initial task statement and the dot plots for grade 7 and grade 10 without revealing the questions that follow. Ask pairs to write down possible mathematical questions that can be answered about the situation. As students share their questions with the class, listen for and amplify questions that ask about the mean weight of the backpacks and whether there is a meaningful difference between the weight of all seventh-grade and tenth-grade backpacks. This will help students write and verbalize their questions about the dot plots of the samples as they attempt to draw conclusions about the populations.
Design Principle(s): Optimize output; Cultivate conversation

Student Task Statement

Consider the question: Do tenth-grade students' backpacks generally weigh more than seventh-grade students' backpacks?

Here are dot plots showing the weights of backpacks for a random sample of students from these two grades:



1. Did any seventh-grade backpacks in this sample weigh more than a tenth-grade backpack?
2. The mean weight of this sample of seventh-grade backpacks is 6.3 pounds. Do you think the mean weight of backpacks for *all* seventh-grade students is exactly 6.3 pounds?
3. The mean weight of this sample of tenth-grade backpacks is 14.8 pounds. Do you think there is a meaningful difference between the weight of all seventh-grade and tenth-grade students' backpacks? Explain or show your reasoning.

Student Response

1. Yes, three seventh-grade backpacks weighed more than the lightest tenth-grade backpack.
2. No. 6.3 pounds is the mean of a sample, and the population mean will probably be at least a little different.
3. Answers vary. Sample response: There is still probably a meaningful difference in the mean weights since there is very little overlap.

Activity Synthesis

The purpose of the discussion is for students to think about how comparing groups by using data from samples differs from comparing groups when the population is known.

Ask partners to share their decision about whether the groups had a meaningful difference with the class.

Consider asking these questions for discussion:

- "Compare the information in this activity to the information about team heights given in an earlier lesson." (In that lesson, we had data from the entire population and here it is only a sample.)
- "Is the overlap of the data more important when you only have a sample or when you have data from the population? Explain your reasoning." (It is more important when you only have a sample. If there is overlap with only some of the data, it's possible there is more overlap when we include more data from the population.)
- "Is it *possible* that the data in the two samples were drawn from population data that is identical?" (It is unlikely, but possible.)

Access for Students with Disabilities

Representation: Internalize Comprehension. Use color and annotations to illustrate student thinking. As students describe their reasoning and the relationships they noticed for each sample, use color and annotations to scribe their thinking on a display of each problem so that it is visible for all students.
Supports accessibility for: Visual-spatial processing; Conceptual processing

18.3 Do They Carry More?

15 minutes
The data in the previous activity came from only one sample for each grade. This may not be enough information to make a very good determination about the entire seventh and tenth grade populations. In this activity, students look at different samples from the same population to see that their means are relatively close based on the MADs of the samples. This concept can be reversed to say that if two samples have means that are *not* very close, then the samples likely came from populations that are quite different. A general rule is given to determine whether two populations are meaningfully different based on the mean and MAD from a sample of each (MP6).

Addressing

- 7.SP.B.3

Instructional Routines

- MLR7: Compare and Connect

Launch

Keep students in groups of 2. Allow students 5 minutes of partner work time, then pause the class to assign samples and explain the general rule.

Ask students to pause after the third question in order to explain the general rule and assign a sample to each group. After all students have paused, assign each group one of the 10 samples to work with for the last 2 questions. Further, explain to students:

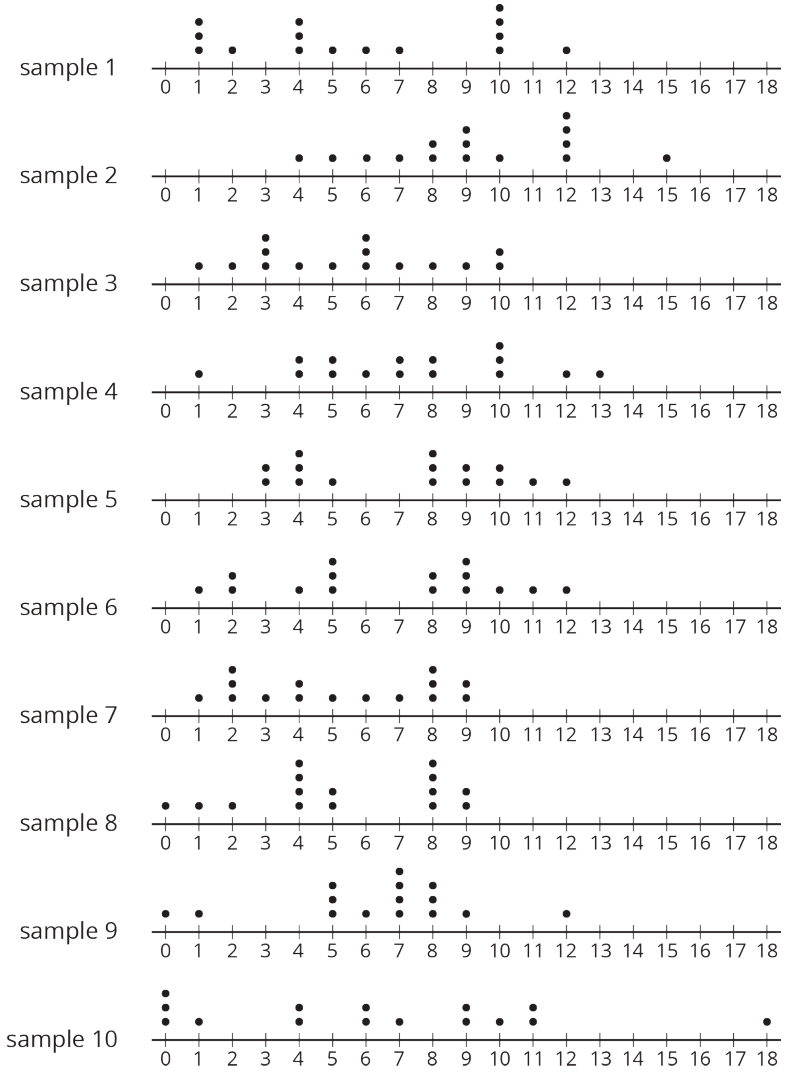
- As a general rule, if two populations have the same mean (or median) and similar variability, the sample means (or medians) should be within 2 MADs (or IQRs) of one another.
- If the sample means (or medians) are more than 2 MADs (or IQRs) apart, it is very likely that the population means (or medians) are different. We will say that there is a *meaningful difference* between the two population means (or medians).
- If the sample means (or medians) are less than or equal to 2 MADs (or IQRs) apart, it is more difficult to say that the two population means (or medians) are very different. In this case we

will say that the samples do not provide evidence that the population means (or medians) differ.

Give students 5 more minutes of partner work time followed by a whole-class discussion.

Student Task Statement

Here are 10 more random samples of seventh-grade students' backpack weights.



sample number	mean weight (pounds)
1	5.8
2	9.2
3	5.5
4	7.3
5	7.2
6	6.6
7	5.2
8	5.3
9	6.3
10	6.4

1. a. Which sample has the highest mean weight?

   b. Which sample has the lowest mean weight?

   c. What is the difference between these two sample means?

2. All of the samples have a mean absolute deviation of about 2.8 pounds. Express the difference between the highest and lowest sample means as a multiple of the MAD.

3. Are these samples very different? Explain or show your reasoning.

4. Remember our sample of tenth-grade students' backpacks had a mean weight of 14.8 pounds. The MAD for this sample is 2.7 pounds. Your teacher will assign you one of the samples of seventh-grade students' backpacks to use.

    a. What is the difference between the sample means for the the tenth-grade students' backpacks and the seventh-grade students' backpacks?

    b. Express the difference between these two sample means as a multiple of the larger of the MADs.

5. Do you think there is a meaningful difference between the weights of all seventh-grade and tenth-grade students' backpacks? Explain or show your reasoning.

Student Response

1. a. Sample 2

    b. Sample 7

    c. 4 pounds ($9.2 - 5.2 = 4$)

2. $4 \approx 2.8 \cdot 1.43$ since $4 \div 2.8 \approx 1.43$

3. No. All of the samples came from the same population and are all within 2 MADs of one another.

4. Answers vary. Possible responses:

sample	1	2	3	4	5	6	7	8	9	10
difference in means	9	5.6	9.3	7.5	7.6	8.2	9.6	9.5	8.5	8.4
multiplier for the MAD	3.2	2	3.3	2.7	2.7	2.9	3.4	3.4	3.0	3.0

5. Yes. The difference in the means is more than 2 MADs, so the means have a meaningful difference.

Activity Synthesis

The purpose of the discussion is for students to understand the general rule for determining if two samples suggest a meaningful difference between their populations.

Select at least one group assigned to each of the samples to share their responses to the last 2 questions and record for all to see. Note that all 10 samples from the seventh-grade students

have means that are within 2 MADs of one another, but the mean from the tenth-grade student sample is at least 2 MADs away from the mean of each of the seventh grade student samples.

Note that the general rule only has two possible outcomes: "There is a meaningful difference." or "There is not enough information to say there is a meaningful difference." If the means are less than 2 MADs apart, the general rule cannot say whether two samples were drawn from populations that contain identical data.

Ask students, "Based only on the dot plots for the 10 samples, would you have guessed that they all might have come from the same population? Explain your reasoning." (Maybe. There is a lot of overlap among all of the samples.)

Access for Students with Disabilities

Representation: Develop Language and Symbols. Create a display of important terms and vocabulary. Invite students to suggest language or diagrams to include that will support their understanding. Include the following term and maintain the display for reference throughout the unit: meaningful difference.
Supports accessibility for: Language; Conceptual processing

Access for English Language Learners

Representing, Speaking, Listening: MLR7 Compare and Connect. Invite students to share their written responses to the last two questions with 2–3 other students. As students investigate each other's work, ask students to make observations about the difference in sample means for the tenth-grade and seventh-grade students' backpacks. Ask students how this observation helps them determine whether there is a meaningful difference between the weights of the backpacks. Listen for and amplify the language students use to describe the general rule for determining whether two samples suggest a meaningful difference between their populations. This will help students make sense of the general rule and discuss how it is used with a variety of samples.
Design Principle(s): Optimize output (for representation); Maximize meta-awareness

18.4 Steel from Different Regions

15 minutes
In previous lessons, students used sample data to estimate population means and proportions and determined if there is a meaningful difference in population means based on sample means. In this activity, students practice using the general rule developed in the previous activity by estimating the measure of center for a population and comparing populations based on those estimates as well as the associated measure of variability (MP3).

Addressing

- 7.SP.B.3
- 7.SP.B.4

Instructional Routines

- MLR1: Stronger and Clearer Each Time

Launch

Keep students in groups of 2.

Explain to students that different regions had different raw materials and techniques for constructing metal. One way of testing ancient metal is by looking at the carbon content in the steel. In some cases, this content could determine the region where the metal was made.

Ask students how the general rule from the previous activity might be adapted to use median and interquartile range (IQR) rather than mean and MAD.

Allow students 10 minutes of partner work time followed by a whole-class discussion.

Access for Students with Disabilities

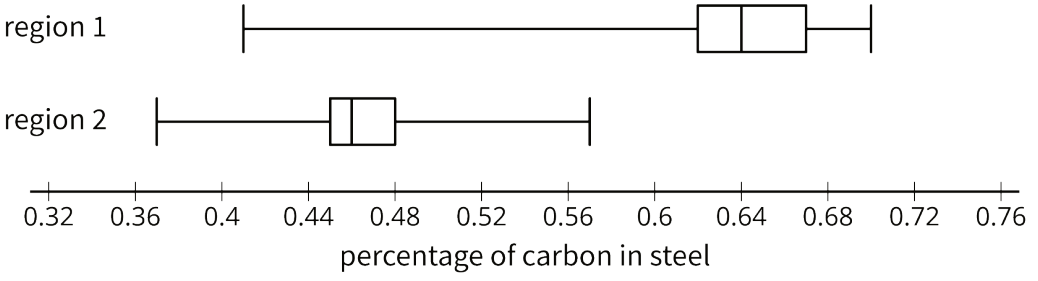
Action and Expression: Internalize Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organizational skills in problem solving. For example, pause to check for understanding after 3–5 minutes of work time.
Supports accessibility for: Organization; Attention

Access for English Language Learners

Writing, Conversing: MLR1 Stronger and Clearer Each Time. Ask students to write a response to the final question: "The anthropologists who conducted the study concluded there was a meaningful difference between the steel from these regions. Do you agree? Explain or show your reasoning." Ask each student to meet with 2–3 other partners for feedback. Provide students with prompts for feedback that will help them strengthen their ideas and clarify their language. For example, "What is the general rule for determining a meaningful difference between populations?" and "How do you know there is a meaningful difference between the steel from these regions?" Students can borrow ideas and language from each partner to refine and clarify their original explanation. This will help students revise and refine their ideas about how to adapt the general rule to use median and IQR rather than mean and MAD.
Design Principle(s): Optimize output (for explanation); Cultivate conversation

Student Task Statement

When anthropologists find steel artifacts, they can test the amount of carbon in the steel to learn about the people that made the artifacts. Here are some box plots showing the percentage of carbon in samples of steel that were found in two different regions:



1. Was there any steel found in region 1 that had:

   a. *more* carbon than some of the steel found in region 2?

   b. *less* carbon than some of the steel found in region 2?

2. Do you think there is a meaningful difference between all the steel artifacts found in regions 1 and 2?

3. Which sample has a distribution that is *not* approximately symmetric?

4. What is the difference between the sample medians for these two regions?

	sample median (%)	IQR (%)
region 1	0.64	0.05
region 2	0.47	0.03

5. Express the difference between these two sample medians as a multiple of the larger interquartile range.

6. The anthropologists who conducted the study concluded that there was a meaningful difference between the steel from these regions. Do you agree? Explain or show your reasoning.

Student Response

1. a. Yes. Most of the steel from region 1 had more carbon in it than steel from region 2.

   b. Yes. Since the left end of the region 1 box plot overlaps with the box plot for region 2, there was at least 1 piece of steel that had less carbon in it than some of the steel from region 2.

2. Answers vary. Sample response: Based on the box plots, there is some overlap, but the boxes look so far apart that I think there will be a meaningful difference.
3. The distribution for region 1 is not symmetric with the very long segment on the left.
4. 0.17%, since $0.64 - 0.47 = 0.17$
5. $0.17 \approx 0.05 \cdot 3.4$
6. I agree with the anthropologists. There is evidence of a meaningful difference because the difference in sample medians is greater than 2 IQRs.

Activity Synthesis

The purpose of the discussion is for students to understand how to adapt the general rule for determining a meaningful difference between populations to median and IQR.

Consider asking these questions for discussion:

- "Why did this problem use median and IQR instead of mean and MAD?" (Since the distribution for region 1 is not symmetric, it makes more sense to use the median. Also the box plots will show the median and IQR, but there is not a good way to know the mean and MAD.)
- "Is there any overlap in the data from the two regions?" (Yes. The smallest percentage of carbon from the region 1 was well below the median from region 2 while the typical percentage of carbon from region 1 is much greater than from region 2.)
- "On the box plot in the activity, draw a dot two IQRs above the median for region 2. Then draw a star two IQRs below the median for region 1. How do these help you see that there is a meaningful difference in the medians?" (The dot is at 0.53 and the star is at 0.54. Since the median for region 1 is not below the dot nor is the median for region 1 above the star, there must be a meaningful difference.)
- "A piece of steel is found in a place between the two regions sampled. Would testing the percentage of carbon from this metal be useful in determining the region from which it came?" (Yes. Since there is a meaningful difference in the percentage of carbon in the steel from the two regions, it should give a good indication which region created the metal.)

Lesson Synthesis

Consider asking these discussion questions to emphasize the main ideas from this lesson:

- "When is it useful to use a median rather than a mean?" (It is useful when the distribution is not approximately symmetric.)
- "What values do you need to calculate from a sample to use the general rule for determining if the measures of center of two populations are meaningfully different?" (The measure of center and measure of variation for each sample should be calculated to compare the groups.)

- "What is the general rule used to determine if the means of two populations are meaningfully different?" (If the difference between the means for the two samples is greater than twice the greater of the MADs, then the means are meaningfully different.)

18.5 Teachers Watching Movies

Cool Down: 5 minutes
This cool-down assesses whether students understand the general rule set out to identify whether the measures of center for two groups are meaningfully different based on a sample of data from each group.

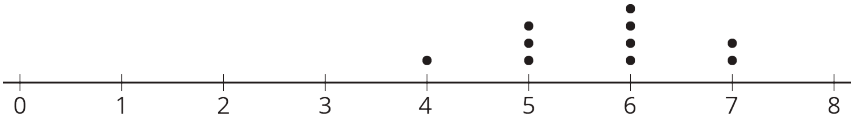
Addressing

- 7.SP.B.3

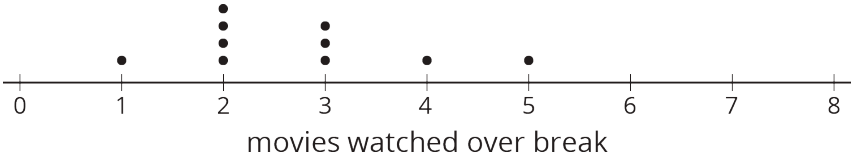
Student Task Statement

Noah is interested in comparing the number of movies watched by students and teachers over the winter break. He takes a random sample of 10 students and 10 teachers and makes a dot plot of their responses.

Students:



Teachers:



Noah then computes the measures of center and variability for each group:

- Students: Mean = 5.7 movies, MAD = 0.76 movies
- Teachers: Mean = 2.7 movies, MAD = 0.9 movies

1. Is Noah's choice of mean and MAD appropriate for the data he has? Explain your reasoning.
2. Should Noah conclude that there is a meaningful difference in the mean number of movies watched over winter break between the two groups? Explain your reasoning.

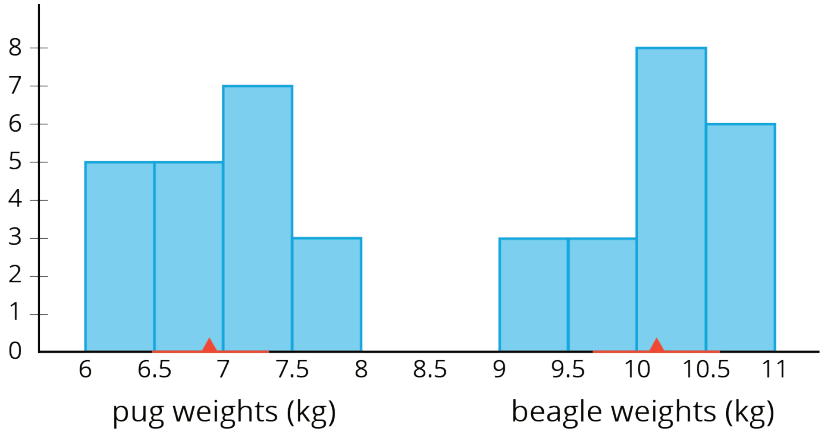
Student Response

1. Yes. Since both samples are approximately symmetric, using the mean is a good choice.

2. Yes. Since the difference in the means is greater than 2 MADs, there is a meaningful difference in the mean number of movies watched. $(5.7 - 2.7) \div 0.9 \approx 3.33$

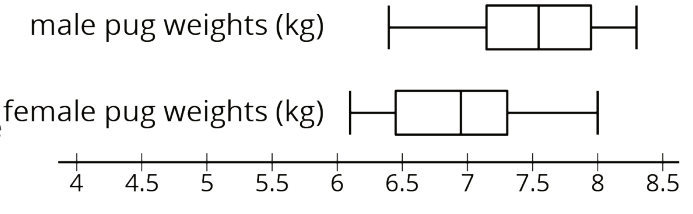
Student Lesson Summary

Sometimes we want to compare two different populations. For example, is there a meaningful difference between the weights of pugs and beagles? Here are histograms showing the weights for a sample of dogs from each of these breeds:



The red triangles show the mean weight of each sample, 6.9 kg for the pugs and 10.1 kg for the beagles. The red lines show the weights that are within 1 MAD of the mean. We can think of these as "typical" weights for the breed. These typical weights do not overlap. In fact, the distance between the means is $10.1 - 6.9$ or 3.2 kg, over 6 times the larger MAD! So we can say there *is* a meaningful difference between the weights of pugs and beagles.

Is there a meaningful difference between the weights of male pugs and female pugs? Here are box plots showing the weights for a sample of male and female pugs:



We can see that the medians are different, but the weights between the first and third quartiles overlap. Based on these samples, we would say there is *not* a meaningful difference between the weights of male pugs and female pugs.

In general, if the measures of center for two samples are at least two measures of variability apart, we say the difference in the measures of center is meaningful. Visually, this means the range of typical values does not overlap. If they are closer, then we don't consider the difference to be meaningful.

Lesson 18 Practice Problems

Problem 1

Statement

Lin wants to know if students in elementary school generally spend more time playing outdoors than students in middle school. She selects a random sample of size 20 from each population of students and asks them how many hours they played outdoors last week. Suppose that the MAD for each of her samples is about 3 hours.

Select **all** pairs of sample means for which Lin could conclude there is a meaningful difference between the two populations.

A. elementary school: 12 hours, middle school: 10 hours

B. elementary school: 14 hours, middle school: 9 hours

C. elementary school: 13 hours, middle school: 6 hours

D. elementary school: 13 hours, middle school: 10 hours

E. elementary school: 7 hours, middle school: 15 hours

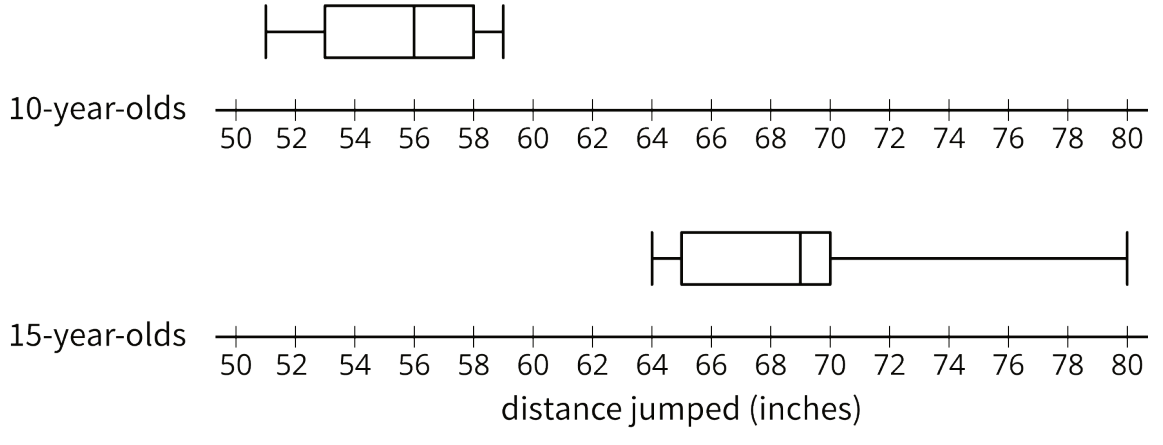
Solution

["C", "E"]

Problem 2

Statement

These two box plots show the distances of a standing jump, in inches, for a random sample of 10-year-olds and a random sample of 15-year-olds.



Is there a meaningful difference in median distance for the two populations? Explain how you know.

Solution

Yes, the difference in medians is 13 inches. This difference is more than 2 IQRs (the IQR is 5 and $13 \div 5 = 2.6$), so there is a meaningful difference in the median distances for 10-year-olds and 15-year-olds.

Problem 3

Statement

The median income for a sample of people from Chicago is about $60,000 and the median income for a sample of people from Kansas City is about $46,000, but researchers have determined there is not a meaningful difference in the medians. Explain why the researchers might be correct.

Solution

The medians differ by $14,000, but if the IQR is larger than about $7,000, there will not be a meaningful difference between the median salaries in the two cities.

Problem 4

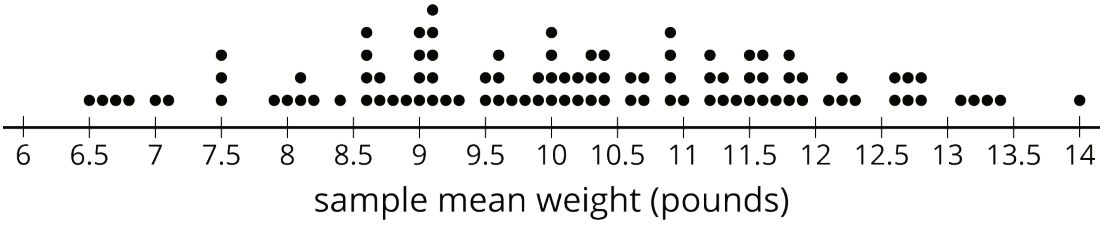
Statement

A farmer grows 5,000 pumpkins each year. The pumpkins are priced according to their weight, so the farmer would like to estimate the mean weight of the pumpkins he grew this year. He randomly selects 8 pumpkins and weighs them. Here are the weights (in pounds) of these pumpkins:

2.9	6.8	7.3	7.7	8.9	10.6	12.3	15.3

a. Estimate the mean weight of the pumpkins the farmer grew.

This dot plot shows the mean weight of 100 samples of eight pumpkins, similar to the one above.



b. What appears to be the mean weight of the 5,000 pumpkins?

c. What does the dot plot of the sample means suggest about how accurate an estimate based on a single sample of 8 pumpkins might be?

d. What do you think the farmer might do to get a more accurate estimate of the population mean?

Solution

a. 8.975 pounds

b. About 10 pounds

c. The sample means ranged from about 6.5 to 14 pounds. If the actual population mean is about 10 pounds, this shows that a sample mean based on a sample of size 8 might not be very close to the actual population value.

d. Use a larger sample size.

(From Unit 8, Lesson 17.)